Do Special Techniques for Surveying Sensitive Topics Provide Valid Measurement?

A Validation Design that Detects False Positives*

Marc Höglinger [1] Andreas Diekmann [2]

[1]University of Bern, Institute of Sociology, *marc.hoeglinger@soz.unibe.ch*

[2]ETH Zurich, Chair of Sociology, *diekmann@soz.gess.ethz.ch*

ESA RN21 Midterm Conference: Data Quality in Quantitative Research
Nicosia, October 14, 2016

* Paper being published as "Uncovering a Blind Spot in Sensitive Question Research: False Positives Undermine the Crosswise-Model RRT." *Political Analysis*, forthcoming

Have you ever provided misleading or incorrect information on your tax return?

- Yes
- No

Did you vote in the 2012 US presidential election?

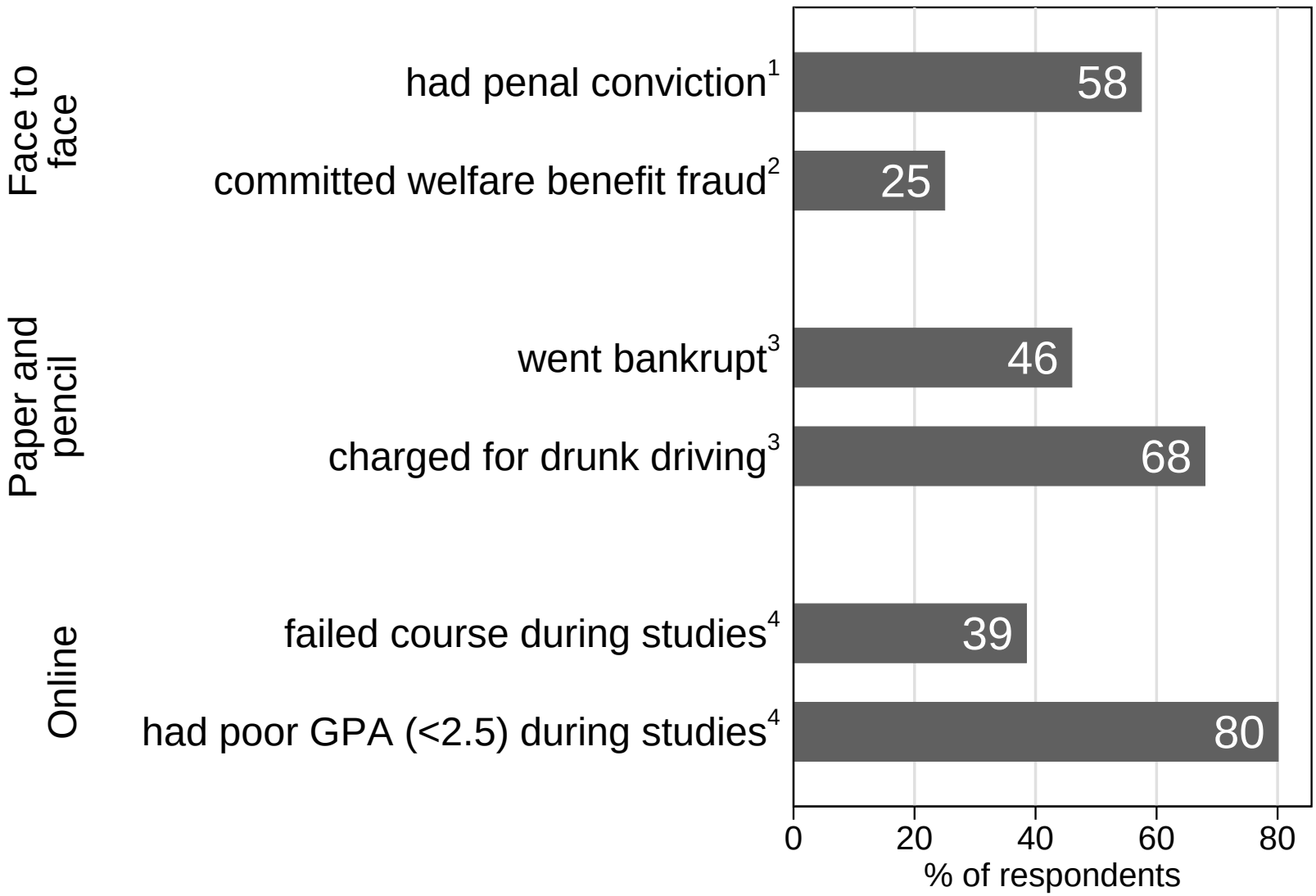
- Yes
- No

Have you ever intentionally taken something from a store without paying for it?

- Yes
- No

Substantial Underreporting of Sensitive Behavior

Proportion of confirmed norm-breakers with truthful self-report (true rate = 100%)



Results from validation studies:

[1] Wolter and Preisendörfer (2013)

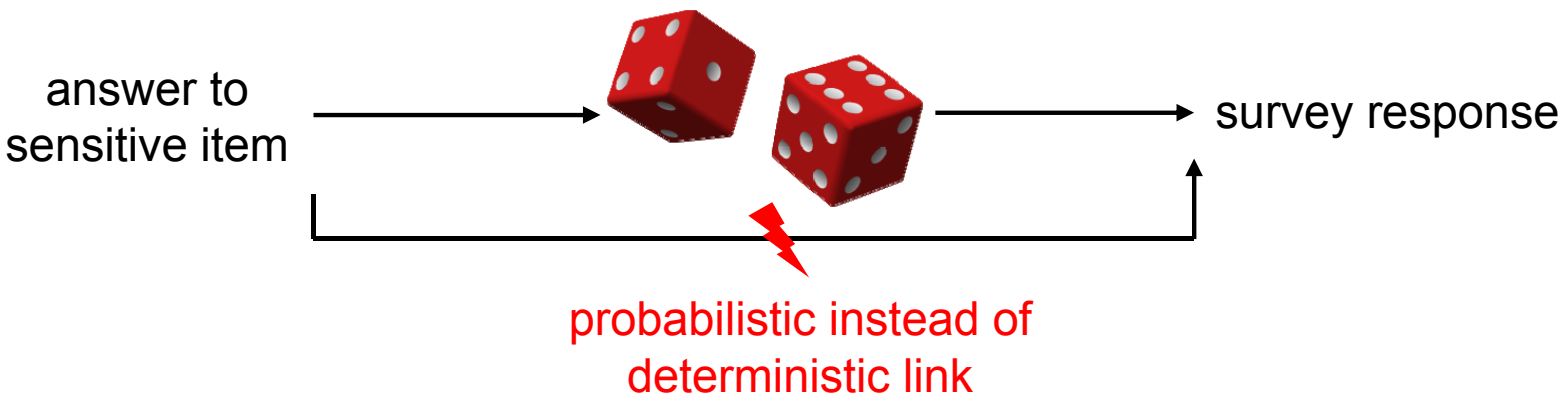
[2] van der Heijden et al. (2000)

[3] Locander, Sudman, and Bradburn (1976)

[4] Kreuter, Presser, and Tourangeau (2008)

The Randomized Response Technique (RRT)

- The RRT (Warner 1965) protects individual's answers with a randomization procedure.
  - random error is introduced in respondents' answers
  - no inference possible from an individual's survey response to her actual answer to the sensitive question
- in turn, respondents should answer (more) honestly



- To analyze RRT data the systematic error is taken into account by adjusting the response variable accordingly.

calculation

The Crosswise-Model RRT (CM)

A recently proposed and seemingly promising new RRT variant (Yu, Tian, and Tang 2008)

Question A:

Is your mother's birthday in January or February?

(If you do not know, please use the birth date of someone else you know.)

Question B:

Have you ever received a donated organ (kidney, heart, part of a lung or liver, pancreas)?

Compare your responses to question A & B. Are they identical or different?

- identical
- different

But, Does it Work? Validation Approaches

- Comparative validation
  - Prevalence estimates are compared under the more-is-better assumption: higher estimates are interpreted as more valid estimates
  - Tenable, if under-reporting, i.e. false negatives, is the only type of misreporting
  - Not tenable, if false positives occur, i.e. if respondents falsely admit sensitive behavior
- Aggregate validation
  - Prevalence estimates are compared to a known aggregate criterion such as official turnout rates (Rosenfeld, Imai, and Shapiro 2015)
  - No DQ as benchmark needed, but also relies on on-sided-lying assumption
- Individual-level validation
  - Self-reports are compared to observed/known behavior or traits at the individual level
  - Preferable, as it can identify false positives as well as false negatives
  - Very difficult to carry out.

CM Judged Favorably in Many Comparative Validations:

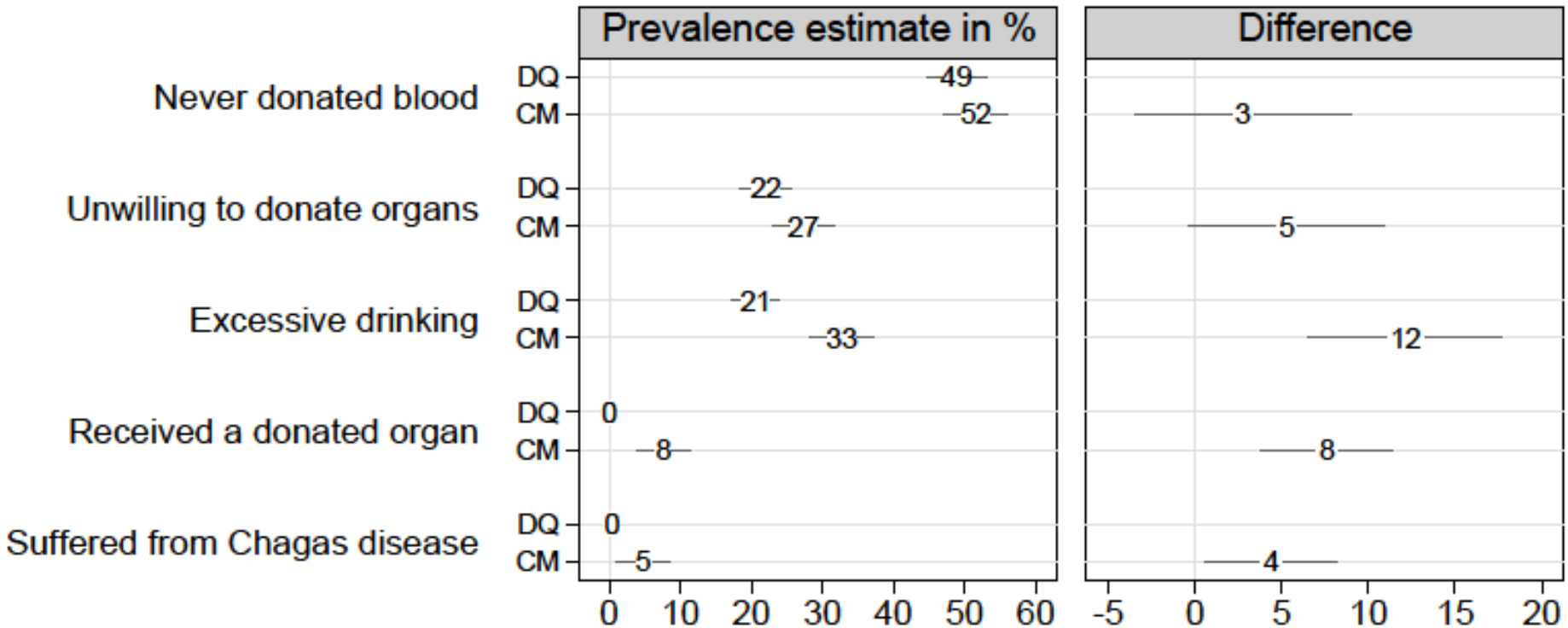
- Adrian Hoffmann and Jochen Musch. 2015. "Assessing the Validity of Two Indirect Questioning Techniques: A Stochastic Lie Detector versus the Crosswise Model". *Behavior Research Methods* (online first)
- Marc Höglinger, Ben Jann, and Andreas Diekmann. 2014. *Sensitive Questions in Online Surveys: An Experimental Evaluation of the Randomized Response Technique and the Crosswise Model*. University of Bern Social Sciences Working Paper No. 9. ETH Zurich and University of Bern. `https://ideas.repec.org/p/bss/wpaper/9.html`
- Ben Jann, Julia Jerke, and Ivar Krumpal. 2012. "Asking Sensitive Questions Using the Crosswise Model. An Experimental Survey Measuring Plagiarism". *Public Opinion Quarterly* 76:32–49
- Martin Korndörfer, Ivar Krumpal, and Stefan C. Schmukle. 2014. "Measuring and Explaining Tax Evasion: Improving Self-Reports Using the Crosswise Model". *Journal of Economic Psychology* 45:18–32
- Mansour Shamsipour et al. 2014. "Estimating the Prevalence of Illicit Drug Use Among Students Using the Crosswise Model". *Substance Use & Misuse* 49:1303–1310
- Adrian Hoffmann et al. 2015. "A Strong Validation of the Crosswise Model Using Experimentally-Induced Cheating Behavior". *Experimental Psychology* 62:403–414
- Daniel W. Gingerich et al. 2015. "When to protect? Using the crosswise model to integrate protected and direct responses in surveys of sensitive behavior". *Political Analysis*: online first

An Enhanced Comparative Validation Design

- Simple design, able to detect systematic false positives without the need of an individual-level criterion.
- Test for false positives with (near) zero-prevalence items:
  - *Have you ever received a donated organ (kidney, heart, part of a lung or liver, pancreas)?*
  - *Have you ever suffered from Chagas disease (Trypanosomiasis)?*
- If a sensitive question technique produces a non-zero estimate $\rightarrow$ false positives, "more-is-better" must be refuted
- Implemented in an online survey on organ donation and health in Germany ($N = 1,685$)

Higher CM Estimates, But More-Is-Better Not Tenable

Crosswise-model produced clearly incorrect estimates for the two zero-prevalence items.



Conclusions

- An up-and-coming implementation of the crosswise-model RRT produced false positives to a non-ignorable extent.
- The crosswise-model's defect could not have been revealed by several previous validations which points to a serious weakness in past research.
- Conclusive assessments of RRT implementations are only possible with validation designs considering false negatives as well as false positives.
- This has also implications for other sensitive question techniques (e.g., Item Count) that so far have been only validated with the same flawed strategies that rely on the "more-is-better" assumption.

Do Special Techniques for Surveying Sensitive Topics Provide Valid Measurement?

A Validation Design that Detects False Positives*

Marc Höglinger [1] Andreas Diekmann [2]

[1]University of Bern, Institute of Sociology, *marc.hoeglinger@soz.unibe.ch*

[2]ETH Zurich, Chair of Sociology, *diekmann@soz.gess.ethz.ch*

ESA RN21 Midterm Conference: Data Quality in Quantitative Research
Nicosia, October 14, 2016

* Paper being published as "Uncovering a Blind Spot in Sensitive Question Research: False Positives Undermine the Crosswise-Model RRT." *Political Analysis*, forthcoming

References I

Gingerich, Daniel W., Virginia Oliveros, Ana Corbacho, and Mauricio Ruiz-Vega. 2015. "When to protect? Using the crosswise model to integrate protected and direct responses in surveys of sensitive behavior". *Political Analysis*: online first.

Hoffmann, Adrian, Birk Diedenhofen, Bruno Verschuere, and Jochen Musch. 2015. "A Strong Validation of the Crosswise Model Using Experimentally-Induced Cheating Behavior". *Experimental Psychology* 62:403–414.

Hoffmann, Adrian, and Jochen Musch. 2015. "Assessing the Validity of Two Indirect Questioning Techniques: A Stochastic Lie Detector versus the Crosswise Model". *Behavior Research Methods* (online first).

Höglinger, Marc, Ben Jann, and Andreas Diekmann. 2014. *Sensitive Questions in Online Surveys: An Experimental Evaluation of the Randomized Response Technique and the Crosswise Model*. University of Bern Social Sciences Working Paper No. 9. ETH Zurich and University of Bern. `https://ideas.repec.org/p/bss/wpaper/9.html`.

Jann, Ben, Julia Jerke, and Ivar Krumpal. 2012. "Asking Sensitive Questions Using the Crosswise Model. An Experimental Survey Measuring Plagiarism". *Public Opinion Quarterly* 76:32–49.

Korndörfer, Martin, Ivar Krumpal, and Stefan C. Schmukle. 2014. "Measuring and Explaining Tax Evasion: Improving Self-Reports Using the Crosswise Model". *Journal of Economic Psychology* 45:18–32.

Kreuter, Frauke, Stanley Presser, and Roger Tourangeau. 2008. "Social Desirability Bias in CATI, IVR, and Web Surveys". *Public Opinion Quarterly* 72:847–865.

Locander, William, Seymour Sudman, and Norman Bradburn. 1976. "An Investigation of Interview Method, Threat and Response Distortion". *Journal of the American Statistical Association* 71:269–275.

Rosenfeld, Bryn, Kosuke Imai, and Jacob N. Shapiro. 2015. "An Empirical Validation Study of Popular Survey Methodologies for Sensitive Questions". *American Journal of Political Science*: (online first).

Shamsipour, Mansour, Masoud Yunesian, Akbar Fotouhi, Ben Jann, Afarin Rahimi-Movaghar, Fariba Asghari, and Ali Asghar Akhlaghi. 2014. "Estimating the Prevalence of Illicit Drug Use Among Students Using the Crosswise Model". *Substance Use & Misuse* 49:1303–1310.

References II

van der Heijden, Peter G. M., Ger van Gils, Jan Bouts, and Joop J. Hox. 2000. "A Comparison of Randomized Response, Computer-Assisted Self-Interview, and Face-to-Face Direct Questioning. Eliciting Sensitive Information in the Context of Welfare and Unemployment Benefit". *Sociological Methods & Research* 28:505–537.

Warner, Stanley L. 1965. "Randomized-response: A survey technique for eliminating evasive answer bias". *Journal of the American Statistical Association* 60:63–69.

Wolter, Felix, and Peter Preisendörfer. 2013. "Asking Sensitive Questions: An Evaluation of the Randomized Response Technique vs. Direct Questioning Using Individual Validation Data". *Sociological Methods & Research* 42:321–353.

Yu, Jun-Wu, Guo-Liang Tian, and Man-Lai Tang. 2008. "Two New Models for Survey Sampling with Sensitive Characteristic: Design and Analysis". *Metrika* 67:251–263.

Analyzing RRT Data

- To analyze RRT data the systematic error is taken into account by using the adjusted response variable $\tilde{Y}$.
- For the crosswise-model:

$$\tilde{Y} = \Pr(S = 1) = \frac{Y + p^{yes,u} - 1}{(2p^{yes,u} - 1)}$$

Y = observed response variable with $Y = 1$ for "identical"
S = actual answer to the sensitive item with $S = 1$ for "yes"
$p^{yes,u}$ = known probability of a "yes" answer to the unrelated question

  - This follows from solving the probability of the response "identical" for $\Pr(S = 1)$

$$\Pr(Y = 1) = \Pr(S = 1) \cdot p^{yes,u} + (1 - \Pr(S = 1)) \cdot (1 - p^{yes,u})$$

Analyzing RRT Data

- To analyze RRT data the systematic error is taken into account by using the adjusted response variable $\tilde{Y}$.
- For the crosswise-model:

$$\tilde{Y} = \Pr(S = 1) = \frac{Y + p^{yes,u} - 1}{(2p^{yes,u} - 1)}$$

Y = observed response variable with $Y = 1$ for "identical"
S = actual answer to the sensitive item with $S = 1$ for "yes"
$p^{yes,u}$ = known probability of a "yes" answer to the unrelated question

  - This follows from solving the probability of the response "identical" for $\Pr(S = 1)$

$$\Pr(Y = 1) = \Pr(S = 1) \cdot p^{yes,u} + (1 - \Pr(S = 1)) \cdot (1 - p^{yes,u})$$

		unrelated question	
		no	yes
sensitive item	no	identical	different
	yes	different	identical

Sensitive Items Surveyed

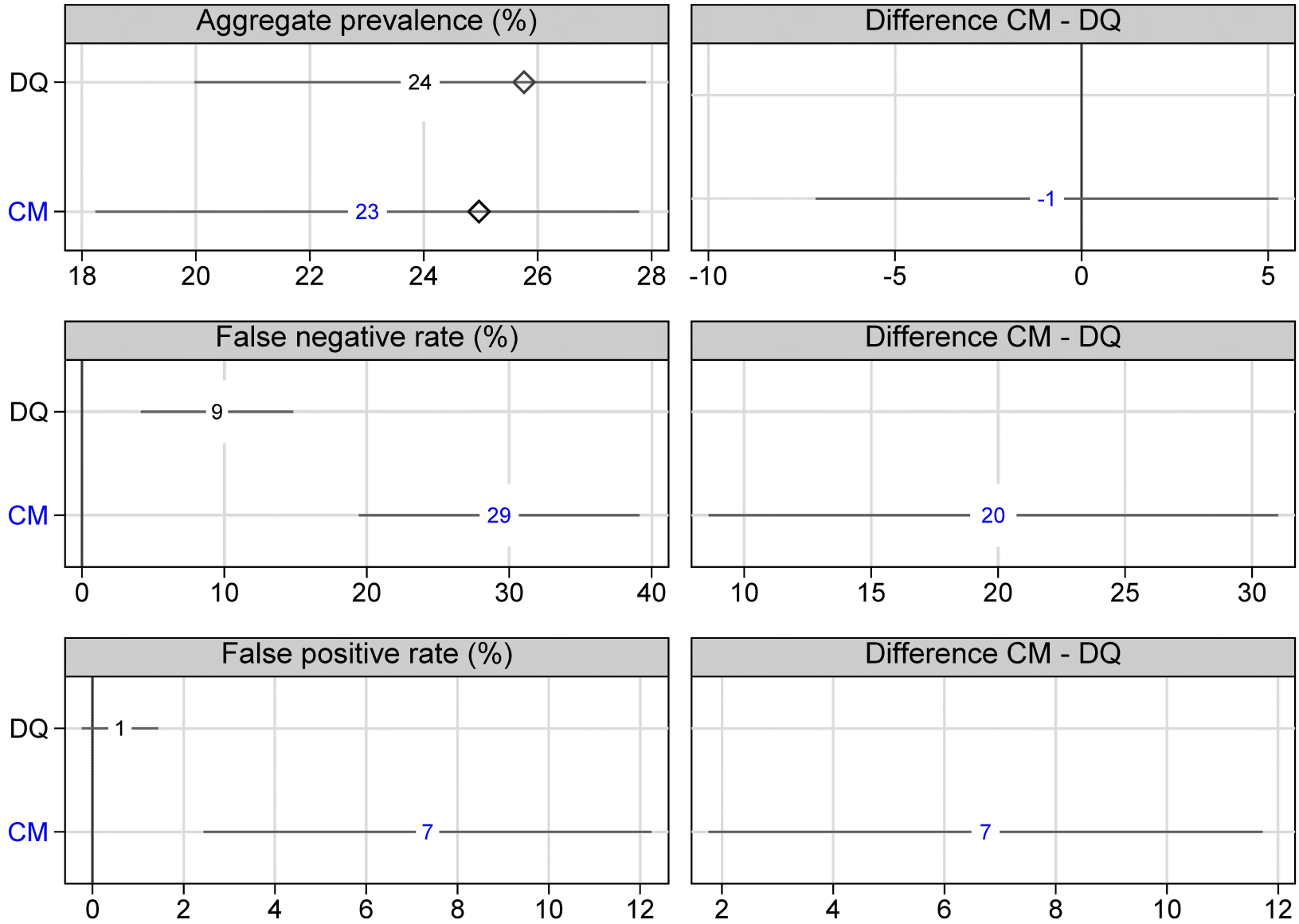
Item	Wording
Copying from other students in exam	In your studies, have you ever copied from other students during an exam?
Using crib notes in exam	In your studies, have you ever used illicit crib notes in an exam (including notes on mobile phones, calculators or similar)?
Taking drugs to enhance exam performance	In your studies, have you ever used prescription drugs to enhance your performance in an exam?
Including plagiarism in paper	In your studies, have you ever handed in a paper containing a passage intentionally adopted from someone else's work without citing the original?
Handing in someone else's paper	In your studies, have you ever had someone else write a large part of a submitted paper for you or have you handed in someone else's paper as your own?

Estimates as displayed in the figure (SE in parenthesis)

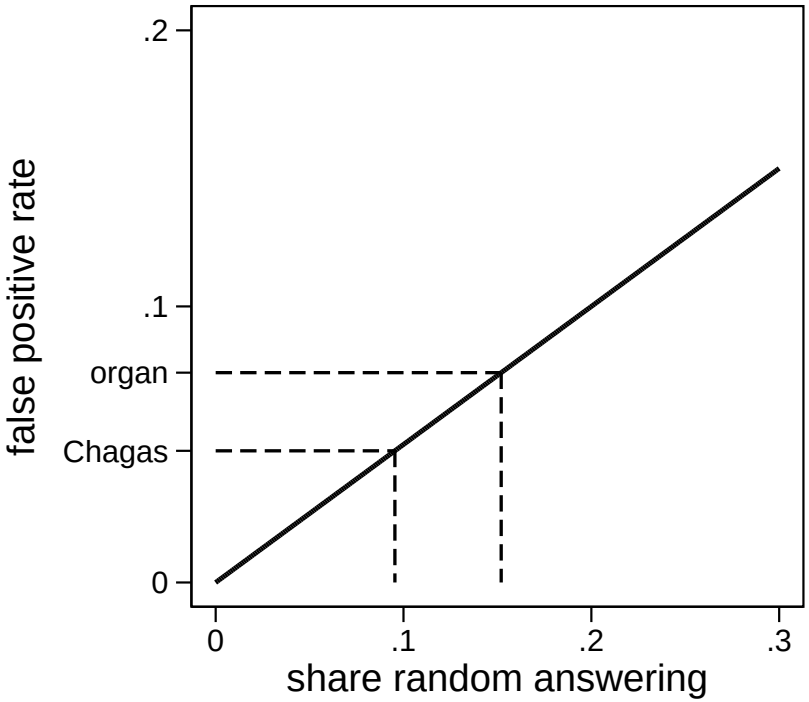
	Never donated blood	Unwilling to donate organs	Excessive drinking	Received a donated organ	Suffered from Chagas disease
Levels					
Direct questioning (DQ)	48.82	22.01	20.58	0.00	0.37
	(2.14)	(1.82)	(1.73)	(.)	(0.26)
Crosswise model (CM)	51.58	27.30	32.71	7.60	4.77
	(2.33)	(2.23)	(2.28)	(1.95)	(1.91)
Difference					
CM – DQ	2.76	5.29	12.13	7.60	4.40
	(3.16)	(2.88)	(2.86)	(1.95)	(1.92)
N	1669	1641	1672	1669	1669

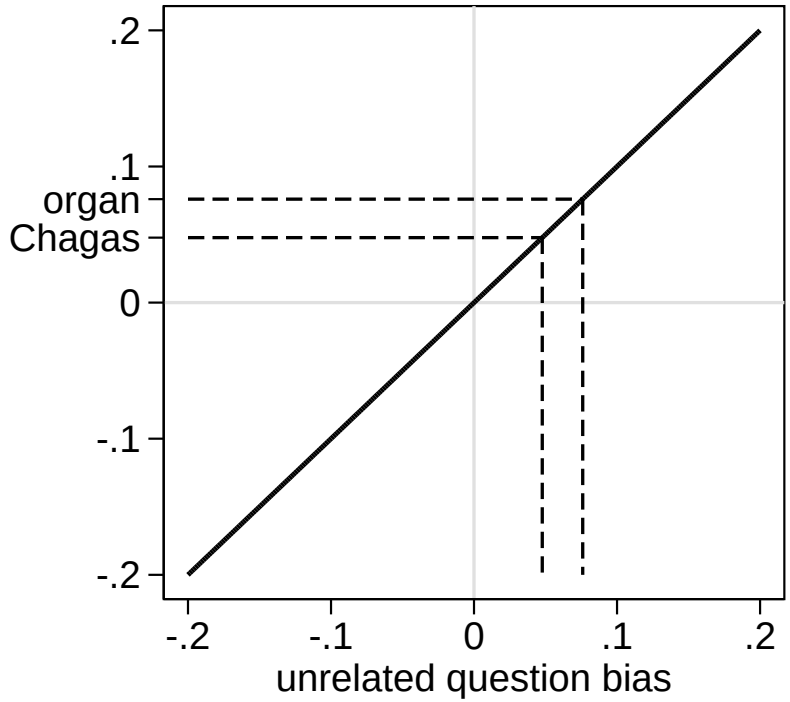
Individual-Level Validation of Abitur-Item

results are corroborated: the crosswise-model implemented produced false positives



Effect of random answering and unrelated question bias on false positive rate for zero-prevalence items





Dashed lines indicate false positive rates found and the corresponding extent of error necessary to generate them.

Notes: With an expected "yes"-probability for the unrelated questions of 0.18 as in the CM implemented. If the "yes"-probability is inverted to 0.82 (half the sample) random answering has the same effect, but the effect of the unrelated question bias goes in the opposite direction.

Exploring Causes of False Positives

Not clearly related to any of our experimental manipulations. Correlates Simulations

Effects of CM implementation details on false positive rate

	Percentage points change	SE
With "don't know" response option	-4.48	(2.79)
Response order different - identical (vs. inverse)	-1.18	(2.79)
Unrelated question on father (vs. mother)	-2.82	(2.87)
Unrelated question on acquaintance (vs. mother)	2.69	(2.91)
Unrelated question on birthday (vs. birth month)	2.04	(2.73)
Yes-probability unrelated question .82 (vs. .18)	-2.10	(2.79)
Item position (linear)	0.09	(0.96)
Item position 1st or 2nd (vs. 4th or 5th)	-1.54	(3.77)

Notes: Bivariate regressions on pooled responses to zero-prevalence items. Standard errors corrected for clustering in respondents. $N = 2,243$. $^{*}p < 0.05$

Exploring Correlates of False Positives

Positively associated with speeding through the CM explanation and with socially desirable responding (MC-scale).

Bivariate associations between respondents' behavior and personal characteristics and false positive rate

	Percentage points change	SE
Among fastest 10% on CM introduction screen	9.05	(4.87)
Among fastest 10% answering sensitive items (without intro)	-4.33	(4.46)
Clicked button referring to RRT Wikipedia link	6.05	(3.90)
Social desirability (Crown-Marlowe scale)	1.62*	(0.80)
Accomplished the university entrance qualification	-5.17	(3.53)
Age	-0.03	(0.10)
Female	-1.73	(2.95)

Notes: Bivariate regression on pooled zero-prevalence items. Standard errors corrected for clustering in respondents. N from 2,208 to 2,243. $^{*}p < 0.05$